(54) Title: TELECOMMUNICATION INTERACTION ANALYSIS
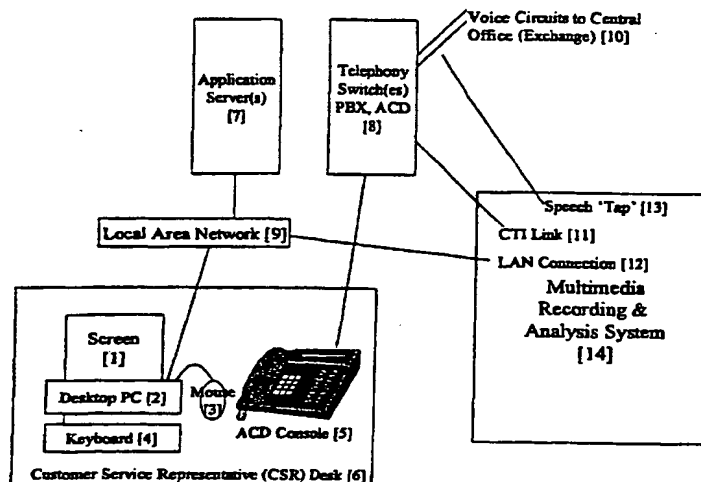
(57) Abstract: The present invention provides for a method of monitoring sets of related communication signal streams comprising the steps of analysing the content or parameters associated with a component of one of the signal streams according to a first analysis criteria, analysing a second component of a related signal stream or parameter associated therewith, according to a second analysis criteria, providing results of the analysis of the said one of the signal streams and which is responsive to the said analysis according to the second criteria. Also, the analysis of the energy envelope representative of at least one communication signal can be provided for and the method of the present invention can further include steps of conducting speech recognition of the identification of words and/or phrases within a communications traffic stream and in which the scale and/or nature of recognition analysis applied to the speech recognition is varied responsive to the analysis of content or parameters associated with the communication stream.

Telecommunication Interaction Analysis

The present invention relates to the analysis of communication signals and in particular such signals
5   representing interaction between the users of a telecommunication system.

Commercial organizations have, for some time, taken the step of recording communications streams such as telephone calls
10 between their staff and their customers. Traditionally this has been necessary to help satisfy regulatory requirements or to help resolve disputes. More recently, the emphasis has moved towards the review of such communications interactions from a quality perspective: the aim being to identify good
15 and bad aspects and characteristics of communication exchanges with a view to improving the level of customer service given.

Also, a record of activity as occurring on an associated
20 display such as a PC screen can also be made and can serve to improve the completeness of a communication-exchange review procedure. In this manner, the reviewer is then able to ascertain how accurately staff are entering information provided during a telephone conversation. Also, particular
25 aspects of an employee's data entry skills and familiarity with the application can be reviewed by recording keystrokes and mouse movement/clicks etc.

So-called Call Detail Recording systems have been employed
30 in order to allow for the prevention of abuse of telephone systems and to apportion costs to the relevant department or individual making the calls. Originally, such records were printed out directly from the Private Automatic Branch Exchange (PABX) onto a line printer. Systems are also now
35 available that are able to store this information in a

database allowing more sophisticated reporting and for the searching of calls on the basis of one or more of the details related to the stored call.

5   Several systems have been developed, for example the AutoQuality system available from e-Talk, and the eQuality system available from Witness Systems and also the present applicant's QualityCall system, that employ call recording in combination with call detail recording and a database

10   application to perform routine monitoring of calls with the intention of identifying weaknesses in the performance of individual Customer Service Representatives (CSRs). Typically a small percentage of the CSRs' calls are reviewed and scored against a set of predetermined criteria to give

15   an indication of the performance of the member of staff.

Also of relevance is the current state of the art of speech recognition systems. First, the automation of simple interactions previously conducted via human interaction, or

20   via touch tone menus, can be achieved. Secondly, dictation products are available that can translate the contents of an audio input into text even though they may exhibit error rates that are greater than would be acceptable if a meaningful transcription of the call was required.

25

Recording systems are also available that can be arranged to provide for the analysis of the content of, for example, a communications stream. Systems providing for the recording of particular events, or incidents, that might arise during

30   a telephone conversation, and the time at which such events or incidents occur within a communications interaction have also been developed.

Such known systems however, and in particular quality-

35   monitoring systems, exhibit disadvantages and limitations

and are discussed in International Application WO 01/52510.

For example, such systems tend to be extremely labour
intensive. The time required to review an interaction can
5   typically take at least as long as the original interaction
lasted. It can also prove necessary to listen to and review
the recording of the interaction several times. For example,
an initial review may be required in order to determine the
content and type of call, and whether or not it is complete
10 enough and appropriate to allow for full evaluation. If so,
it is then re-played completely for review against pre-
determined scoring criteria. It then has to be re-played
again for review with the CSR who took the call.

15 Known systems also prove unable to identify infrequent
problems. Because of the time taken to review a call, it is
rare that more than a fraction of one percent of all calls
are evaluated and reviewed. This renders the reviewed calls
statistically very poor for identifying rare problems.
20 Realistically, such systems can only hope to provide an
indication of the average quality of interactions carried
out by each CSR.

Increasingly, CSRs are expected to be multi-skilled and to
25 handle a wide range of different types of calls. Unless many
more hours are spent reviewing calls, it is impossible
effectively to identify problems that occur in a small
proportion of a CSR's calls. If problems are only rarely
spotted, it then becomes very difficult to recognize
30 underlying patterns since such instances become isolated.

Also, such known systems are very much subjective and, even
with the best training and call-evaluation coaching, the
evaluator will apply at least some degree of subjectivity to
35 their evaluation particularly with softer aspects of

assessment such as customer satisfaction levels. While such
systems can provide tools that highlight discrepancies
between different evaluators, they are restricted in that
they cannot serve to prevent such subjectivity.

5

Known systems also are not generally normalized.   For
example, the manner in which organizations choose to measure
call quality is entirely at their discretion and so a 95%
quality rating achieved by one organization may in reality
10   be worse than the 90% rating achieved by another
organization employing a stricter marking schema. This lack
of consistency between organizations makes it difficult, for
example, for organizations to evaluate how they compare with
their industry peers or indeed with other industries.

15

The present invention seeks to provide for a system and
related method which can offer advantages over known such
systems and methods.

20   According to an aspect of the present invention there is
provided   a   method   of   monitoring   sets   of   related
communication   signal   streams   comprising   the   steps   of
analysing   the   content   or   parameters   associated   with   a
component of one of the signal streams according to a first
25   analysis criteria;
analysing a second component of a related signal stream or
parameter   associated   therewith,   according   to   a   second
analysis criteria;
providing results of the analysis of the said one of the
30   signal streams and which is responsive to the said analysis
according to the second criteria.

This   aspect   of   the   present   invention   therefore
advantageously provides for the linking of the analysis of
35   related data streams so as to enhance the analysis of at

least one of the streams.

Advantageously, the said first analysis criteria is arranged
to be selected by means of the said second criteria.

5

Also, the said analysis of the content or parameters
associated therewith and the analysis of the signal stream
are combined to provide a composite output parameter.

10 Further, the analysis according to the second criteria
occurs prior to the analysis according to the said first
criteria.

According to another aspect of the present invention there
15 is provided a communication monitoring system including
means for determining an energy envelope representative of
at least one communication signal, and means for providing
for the subsequent analysis of the said energy envelope.

20 The monitoring and analysis of the energy envelope
represents a particularly efficient and accurate means for
determining a variety of aspects and characteristics of, for
example, a two-way telephone conversation.

25 Advantageously, at least two energy envelope files are
employed and this can serve to allow for the advantages of
stereo recording without disadvantageously doubling storage
requirements.

30 Appropriately, the system can be arranged to allow for the
selective analysis of the energy envelope and, in
particular, analysis of the energy envelope representative
of the final section(s) of, for example, a telephone
call/conversation.

35

6

Further, the energy envelope may be analyzed so as to identify clipping of the signal which can be indicative of periods of raised voices, or shouting, within the communications traffic stream.

5

Also, talk/silence ratios can advantageously be determined from the energy envelope so as to identify periods when no communication signals arise, for example, during music-on-hold periods or when a ring-tone is being generated.

10

The system advantageously further includes storage means for storing the energy envelope for subsequent analysis.

Also, the pattern of activity towards the very end of a call
15 can give indications of abnormal termination - calls being cut-off in the middle of speech or where there is no activity from one or other party for several seconds prior to the end of the call.

20 A further indication of interest is the speed and clarity of speaking which can be inferred from the gaps between utterances and the average duration of each word spoken.

This aspect of the present invention also advantageously
25 provides for a method of monitoring communication signal including the step of determining an energy envelope representative of at least one communication signal, and subsequently analyzing the said energy envelope. The method can advantageously be conducted in accordance with the
30 system such as that defined above.

According to another aspect of the present invention, there is provided a communications monitoring system including speech recognition means for the identification of words
35 and/or phrases within a communications traffic stream, and

means for varying the scale and/or nature of recognition analysis applied by the speech recognition means.

Advantageously the scale and/or nature of the recognition
5   analysis is arranged to be varied responsive to the identification of at least one party to the communication session. As well as a variety of alternatives, the scale and/or nature of the recognition analysis can be varied on the basis of the length and/or stage of the communication
10   session.

Preferably, the system is arranged to provide speech recognition serving to offer an indication of the level of customer satisfaction. Means can also be provided to
15   generate a score signal indicative of such a level of satisfaction.

Advantageously, separate storage means are provided for storing positive and negative scores.
20
The system advantageously can also include means for monitoring the operation of a user interface device, the output of which can advantageously be employed in controlling the scale and/or nature of the recognition
25   analysis. For example, since a particular area of interest to a customer can, at any time, be indicated by means of information displayed at a graphical display device associated with the system, a speech recognition module can be operated in a then predetermined manner having regard to
30   the topic being discussed, and thus the keywords and words likely to be spoken.

This further aspect of the present invention also advantageously provides for a method of monitoring a
35   communications traffic stream including the application of

speech recognition to an audio signal arising from part of that communication and having a varying scale and/or nature of recognition analysis and, advantageously, in accordance with the system as defined above.

5

According to a further aspect of the present invention, there is provided a communications monitoring system including a user interface device for allowing manual input of data to the system, and uses interactions with the said

10 user interface device.

The manner and nature of use of any such user interface device can advantageously provide further information which can usefully be employed in assisting with the monitoring

15 and analysis of the communications traffic stream.

The system can also advantageously allow for monitoring of the accuracy with which a user employs such a user interface device by, for example, monitoring the use of the backspace

20 or delete key of a keyboard etc.

Also, the use of predetermined features of the user interface device can advantageously serve to delineate different sections of the record of use of the user

25 interface device so as to advantageously associate such different sections with respective different sections of the communications traffic stream.

Also, the joint monitoring of the use of the user interface

30 device and the level and/or nature of communications traffic arising can advantageously serve to identify any potential short comings in the skills/efficiencies of the user, for example, from the analysis of, or relation between any pulses arising in the audio signal and corresponding

35 activity noted at the user interface device.

This further aspect of the present invention also provides for a method of monitoring communications signals including the step of monitoring the use of a user interface device associated with the monitoring system.

5

Advantageously, the invention can also provide for a combination of any one or more of the above-mentioned aspects.

10 The invention can prove advantageous in at least partially automating the assessment and categorization of recordings. That is, by recording and subsequently analyzing various aspects of the interactions, the system automates the measurement of a range of attributes which previously could 15 only be determined by listening/viewing recordings. For example, these can include, customer satisfaction , call structure (ratio of talking to listening), degree of interruption occurring, degree of feedback given to/from customer, CSR's typing speed and accuracy, CSR's familiarity 20 with and use of the computer application(s) provided, training needs, use of abusive language, occurrence of shouting/heated exchanges, degree of confusion, adherence to script, avoidance of banned words/phrases and the likelihood of customer/CSR having been hung up on.

25
As a further advantage, the invention can highlight unusual calls for efficient manual review. By measuring the attributes described above, the calls with the highest/lowest scores on each or a combination of such 30 categories can be presented for review. In addition to having automatically selected the calls most likely to be of interest, the present invention provides for mechanisms to present the candidate calls for efficient review. It does this by retaining information, specifically the start and 35 end times related to incidents within the call that led to

it being selected. By way of such information, the most appropriate parts of the call can be selectively played without human intervention. For example, when reviewing potentially abusive calls, only the audio signals

5 approximate to the point where a swear word was identified need be played in order for the reviewer to determine whether or not this is a genuine case of an abusive call or whether a false indication has occurred. Similarly when reviewing calls that are identified as having terminated by

10 one party hanging-up only the last few seconds of the call need be played i.e. the section just prior to the termination. This can therefore allow rapid unattended replay of successive examples without the user having to interact with the system except perhaps to interrupt

15 operation when a potentially interesting call is heard.

The invention can also advantageously offer an objective analysis of the calls. By applying fixed rules and algorithms to the identification of incidents within the

20 calls, and the subsequent categorization or scoring of calls against predetermined criteria and weighting, the scores derived for a given call are deterministic and consistent. Whilst in some respects, the automated scores may not seem as accurate as could be achieved by a well trained human

25 scorer, the fact that the scores can be determined from a much larger sample of, and ideally all available calls, makes them much less subject to random fluctuations than would occur with the small samples such as are scored manually.

30

Also, the invention can advantageously achieve consistency of analysis. Some aspects of calls that can be measured are independent of the particular products, services or organizations that a customer is dealing with in the

35 interaction. For example, customer satisfaction, if measured

by analysis of the words and phrases spoken by the customer during the call can legitimately be compared across a wide range of organizations and industries. As long as the algorithm used to determine the customer satisfaction rating

5 is kept constant, relative levels of satisfaction can thereby be measured across peer groups and across different industries.

The invention is described further hereinafter, by way of

10 example only, with reference to the accompanying drawings in which:

Fig. 1 is a schematic block diagram of an analysis system embodying the present invention;

15

Fig. 2 is a schematic block diagram of the recording and analysis sub-system of Fig. 1;

Fig. 3 is a schematic diagram illustrating the separating of

20 an embodiment of the present invention;

Fig. 4 is a schematic diagram illustrating a generic analysis module according to an embodiment of the present invention;

25

Fig. 5 is a schematic representation of one particular embodiment of analysis module of the present invention;

Fig. 6A and 6B illustrate graphical displays desirable from

30 an analysis module such as that of Fig. 5;

Fig. 7 is a schematic representation of a further embodiment of analysis module embodying the present invention;

35 Fig. 8 is a schematic representation of yet another

embodiment of analysis module of the present invention;

Fig. 9 is a schematic representation of a further embodiment of analysis module of the present invention; and

Fig. 10 is a schematic representation of yet a further embodiment of analysis module of the present invention.

Turning first to Fig. 1, there is illustrated a multimedia recording and analysis system 14 which is arranged to incorporate the specific methods and systems of the present invention and which is used to monitor the interaction between a CSR and the people/customers and/or systems with whom/which the CSR interacts. Such interaction is conducted by means of a telephone system typically including, as in this example, a console 5 on the CSR's desk and a central switch 8 through which connectivity to the public switched telephone network (PSTN) is achieved via one or more voice circuits 10.

The CSR will typically utilize one or more computer applications accessed through a terminal or PC at their desk 2 with which they can interact by means of the screen 1, mouse 3 and a keyboard 4. The software applications employed may run locally on such a PC or centrally on one or more application server(s) 7.

The system 14 embodying aspects of the present invention can advantageously offer connections to so as to monitor and/or record any required combination of aspects such as the audio content of telephone conversations undertaken by the CSR's by means of a speech tap 13 or the contents of the screen on the CSR's desk during interactions. This latter aspect can require software to be loaded onto the CSR's PC 2 to obtain such information and pass it to the recording system.

The keystrokes, mouse movements and actions of the CSR can be monitored and can also typically require software to be loaded onto the CSR's PC 2. Monitoring of the context of, and data entered into or displayed by, the applications being used can typically require system integration to have the applications pass such information to the recording system. Further, the details of calls placed, queued and transferred etc by the telephony system can be monitored.

It should of course be appreciated that this is merely a typical example of a system that can employ the present invention and numerous variants on this theme are well known, such as the use of Voice Over IP (VoIP), the tapping of the audio signals at the console rather than the trunks, and the use of silent monitoring features to allow for tapping into selected consoles.

Fig. 2 represents a high level view of the major components within a recording and analysis system embodying the present invention. It should be appreciated that such systems are typically deployed across multiple sites and are implemented on multiple computer platforms. However, the major functional blocks remain the same.

Incoming data streams, such as voice, screen content, events etc. 17, and recording control information such as CTI information from an ACD that trigger commands from a desktop application etc. 18 are processed by one or more record processors 19. The net results of such processing are, first the storage of call content in some form of non-volatile storage system such as a disk file storage system 16 can be achieved, and secondly details about the recordings made can be stored in a relational database 15 allowing subsequent search and retrieval on the basis of a number of criteria.

These recordings and the details about them are then subsequently available for one or more search and replay applications that allow users, or other applications such as a customer relationship management system, to access the
5 particular recording(s) they require. One such application is a quality assessment application which will typically make random or pre-programmed selections of recordings and present these to a reviewer for evaluation and subsequent analysis of the results of the said evaluations. Such call-
10 flow recording is described in the present applicant's International application WO 01/52510.

The enhancements to such recording and analysis systems that can be achieved by the present invention relate to methods
15 that act upon the content of the recordings and/or the details about such recordings. In a system of the type shown in Fig. 2, such processes may be advantageously applied at one or more of a variety of points in the system. The optimum location for each method will depend generally on
20 the analysis being performed and also the accuracy required and the topology of the system.

Examples of the options for deploying such methods as described are shown in Fig. 3 and are as follows. First,
25 the point 20 in the system at which the method is employed can comprise part of the record process, with access to the raw, unprocessed, information as received. This may prove to be the only way in which to influence the operation of the recording system as a result of the analyses performed in
30 real-time. Such a location may also be the only point at which unadulterated information is available, for example un-compressed audio that is only stored subsequently to disk once it has been compressed and/or mixed with information/data from other input channels. An alternative
35 location comprises the point at which data is written to

disk. This can prove particularly useful if only a subset of input data is to be recorded. By applying the required algorithms to data at this point, resources are not wasted in attempting to process data that was not actually

5   recorded.

The present applicant's European patent application EP-A-0 833 489 discloses features such as those described above.

10 A further option comprises a location 22 forming part of an offline process. Here, although the overhead of having to query the database and/or read the recording content from disk is incurred, this does allow ongoing 24 hour analysis since it may not prove possible to keep up with the rate of

15 recordings made during the busiest periods of the day.

Advantageously, it can be arranged that such analysis modules are deployed on the CSR'S desktop PCS during periods when the PCS would otherwise be idle. This allows economic

20 deployment of complex analysis such as full speech recognition which, otherwise, would disadvantageously require additional investment in additional processors or would have to be restricted to a much smaller subset of the total recordings.

25

Also, at location 23, some of the analyses may be performed as part of search and replay applications. This is particularly advantageous for analyses that can be performed rapidly on a small set of calls that are already known to be

30 of interest to the application/user in question. The details about the recordings, and the recordings themselves, will in some instances, already have been retrieved by that application and so will be accessible to the analysis tools of the present invention.

35

Such an arrangement is illustrated by reference to Fig. 4 in which call recording details 25 and call recording content 26 are input to an analysis module 24.

5 The analysis algorithm within the module 24 operates on these inputs to produce further details 27 about the recordings and/or further recordings 28 derived from the input recordings.

10 Again, according to deployment, these outputs may simply be used by the application holding the module such as at location 23 or may be written back to the database 15 and/or file system 16. In the latter case it should be noted that the outputs from any module are therefore available as

15 inputs to other modules allowing cascading of analyses such that some modules may produce interim results whilst others further process the outputs or combine them with the outputs of still further modules to produce composite and derived outputs.

20

An example of such a module is shown in Fig. 5 and is arranged to produce an output file for each input voice recording which summaries the audio level or energy present throughout the recording. In its simplest form such an

25 energy envelope module 33 is arranged to operate on an incoming audio signal 30 and convert it to a signed linear encoding scheme 31 if it is not already in such a format. It then averages the absolute value (or, optionally the square of the value) over a fixed interval in the order of

30 typically 50ms. This interval is chosen so that when displayed graphically, the resolution of the samples is sufficient to allow easy visibility of the words and pauses in the recording. An example of a graphical output derived from such an energy envelope' file is shown in Figure 6A

35

These files prove useful in serving as thumbnail graphical overviews of calls as well as serving as useful input for subsequent analysis stages. The energy files avoid the need to retrieve the entire audio recording, and to decompress
5 it, and so make many subsequent analyses viable that would otherwise prove prohibitive due to network bandwidth and/or processing requirements.

As with all other parameters recorded in the invention, the
10 storage may be beneficially accomplished by writing the information in the form of an XML file. The structure of the energy envelope file can be very simple for example it can comprise merely a succession of the average energy values. Beneficially however, the maximum energy value
15 encountered within the file is noted at the start of the file. This allows an application using this file to perform scaling on the file without first having to read the entire file in search of the maximum value.

20 This maximum value is noted by a statistical analysis function 36 as illustrated in Fig. 5 as the recording is being processed. Additional statistics derived from the energy values may also be derived at this time. For example, the ratio of quiet periods (when energy is below a specified
25 threshold for a high proportion of the samples) to active periods can be obtained. Also, the prevalence and location within the call of any periods of clipping, i.e. where the audio signal saturates at the extreme of the available audio range leading to distortion can be identified. This may
30 indicate extreme volume levels such as those arising due to the customer shouting.

This module is advantageously deployed where the audio signal has not yet been compressed. It is much more
35 economical to convert standard telephony signals (e.g. in

G.711 mu-law or A-law) to linear than it is to decompress a
heavily compressed audio signal.

Furthermore, some parameters such as the "degree of
5 clipping" are adversely affected by the compression
algorithms employed.

The module 33 can further advantageously be deployed prior
to any mixing of audio signals such as occurs at location 20
10 in Fig. 3 such that the output energy envelope file reflects
the audio levels in a single direction of speech received.
Thus, although the original transmit and receive signals may
be subsequently mixed into a single audio file for more
efficient storage, the two energy envelope files may be used
15 to produce a clear graphical display as shown in Figure 6B
that highlights who was talking and when, and also enables
interruptions to be highlighted as indicated by arrows A.

Referring now to Fig. 7, an energy envelope analysis module
20 37 takes as its input, one or more energy envelope files 38
plus details about the calls 40 to which they relate.
Typically, the module 37 will serve to analyze the two
energy envelope files relating to the transmit and receive
audio paths for a single call but may also compare a set of
25 energy envelope files for a set of supposedly similar calls.
Statistical analysis indicated at 39, 41 of the input energy
envelope files can be performed to derive output information
such as those discussed as follows.

30 The proportion of talk periods to listen periods within the
call. The frequency of confirmatory feedback from each party
in the call, i.e. when one party is speaking, the   other
will normally respond with an 'uh-huh' or similar utterances
which shows as a brief burst of energy on one channel in the
35 midst of a sustained burst of energy due to the sentence

being spoken on the other. The frequency and proportion of argumentative interruption which can be defined as sustained activity on both channels concurrently for a period exceeding the normal time needed for one party to
5 concede control of the conversation to the other. The proportion of silent periods within the call. The locations of sustained silences within the call and also which party eventually breaks the silence. An unusual call termination pattern different from the usual pattern at the end of a
10 call, when each party speaks briefly, to say goodbye etc. followed by a brief pause and then a loud click as the call is terminated. An abrupt termination of a call within a sustained period of activity by one or other party which can indicate a likely abnormal call-termination. Episodes of
15 shouting or increasing volume, in which the average volume of one or both speakers alters during the course of the call and which can be flagged as a possible indication of a heated conversation.

20 Any of the aforementioned may be combined with a weighting profile that influences the effect of each function of time throughout the call. For example, the determined value of output information preferred talk-to-listen profile may be 50:50 during the first 30% of the call but then may change
25 to a ratio 30:70 thereafter.

A more sophisticated analysis can be performed by utilizing speech recognition tools in order to identify keywords within recordings or to perform large vocabulary
30 transcription of the audio into text. Fig. 8 illustrates such a module 43. The audio streams and, optionally, energy envelope files previously generated 44 are used as an input, along with any pre-existing details 45 about the recordings. The input may initially be sliced at location 46 using the
35 energy envelope files and other details to determine which

portions, if not all, of the recording are to be analyzed by a speech recognition engine 47. The recognition engine 47 is then delivered to a database 49 of entries listing transcript and/or individual words recognized via a database
5 and/or to a file 50 holding similar details directly on the file storage system.

The output from such a speech recognition module 43 is typically one or more of a so-called best guess
10 transcription of the call, or a sequence of recognized words or phrases, their locations within the call and some measure of the likely degree of confidence in their recognition

Such details can be stored for direct searching so as, for
15 example, to find all calls containing a specific word or for further analysis.

The speech recognition module 43 is advantageously deployed where the audio signal has not yet been compressed.
20 Recognition accuracy and the ease of computation are found to be better for an un-compressed signal than for a compressed one.

The speech recognition module 43 is further advantageously
25 deployed prior to any mixing of the audio signals such as at location 20 so that a single speaker can be recognized at a time. This allows the optional deployment of speaker specific recognition models where the speaker is known from the recording details and also ensures that the output is
30 unambiguously linked to the appropriate party to the call. Hence the output is both more accurate and more useful.

Advantageously, if the unmixed stereo recording is unavailable, the speech recognition module 43 may take as
35 its inputs, the mixed audio recording and the energy

envelope files previously generated. These advantageously allow the recognition engine to :-

a) determine which of the two speakers on the call is active at any time and hence apply the most appropriate
5 speaker and vocabulary model enhancing accuracy;

b) label the output with a clear indication as to which party uttered the words detected; and

c) identify more clearly the start and end of utterances which otherwise may merge into one and hence result in lower
10 recognition accuracy as the recognition engine expects a single phrase or sentence in each contiguous utterance rather than two sentences.

Advantageously, the recognition engine 47 is instructed
15 to recognize less than the entire call. As recognition is extremely processor intensive, it can prove beneficial to analyze selected portions of the call. For example, the first 30 seconds can be analyzed to determine the type of call, and the last 30 seconds analyzed to determine the
20 outcome of the call and level of customer satisfaction.

Further, the above partial analysis of the call may be optimized by using the previously derived energy envelope files. Using these energy envelope files, the location and
25 duration of the first and last n utterances by each party can easily be determined and the recognition engine directed to process only these portions of the call. For example, by analyzing the last utterance made by each party it is normally possible to determine the appropriateness of the
30 closure of the call and hence to identify those in which an unusual call closure occurred such as when a CSR hung-up on a customer.

Advantageously, the speech recognition module 43 may only be
35 instructed to analyze a subset of calls that have already

proved to be of potential interest due to any combination of recording details and details derived for example from prior energy envelope analysis.

5   The speech recognition module 43 may also make use of a call flow recording input that indicates the current context of the application(s) being used by the CSR. A vocabulary and grammar model applied by the recognition engine can be influenced by a determination of which application/form and

10 field is active on the CSR's screen. This leads to more accurate recognition and the context can be recorded along with the transcript output allowing subsequent modules to search for words uttered at specific points in the structure interaction flow.

15

Turning now to Fig. 9, there is illustrated a language analysis module 51 which is used to process the output of the speech recognition module. By comparing the words identified in calls 53 against a list of phrases that are of

20 interest, the call can be annotated with database entries 57 and/or additional recorded file information 58 that highlight the presence or absence of these phrases. The output typically includes the start position of the phrase, its duration and confidence of recognition allowing

25 subsequent review of exactly this portion of the call.

Advantageously, the phrases being sought may include wildcard words, e.g. for example the phrase "you've been? helpful" would match a phrase that contained any word

30 between "been" and "helpful".

The phrases can be grouped according to the type of information they indicate. For example, the above phrase would be a customer satisfaction indicator phrase, whereas

35 the "I'm not sure" would be identified as a training need

indicator etc.

Further, each phrase can be allotted a score as to how relevant it is to the type of information sought. For
5 example a simple "thank you" could score +0.1 on the customer satisfaction indicator category whereas "thank you very much" would score +0.3. By storing these relative scores the reviewer can see the relative importance of each phrase matched when reviewing calls.

10
Advantageously, any cumulative score achieved by each call on each of the categories is summed by a score accumulator 55 and the net results for each call are stored to the database 57 and/or the file 58. The score accumulator may
15 apply a time function that weights the scores for specific categories according to the time within the call, whether absolute or relative, that the phrase is recognized. For example, the customer satisfaction indicator would be weighted more heavily towards the end of the call rather
20 than the beginning as the customer may already be happy or upset due to other factors at the start of the call. The success of the call is more accurately determined by the customer's state at the end of the call.

25 In situations where both positive and negative scores are assigned to phrases in the same category the system is arranged to separate total positive and negative scores, rather than merely seeking to cancel these out. A call with extremes of positive and negative satisfaction is naturally
30 of more interest and different from one where no expression of satisfaction is made.

Advantageously, the language analysis module 51 may also make use of Call Flow Recording input that indicates the
35 current context of the application(s) being used by the CSR.

Phrases and their scores can be linked to specific contexts within the application and their scores and applicability varied according to this context.

5 The module 51 may further, also serve to operate on the output of a keystroke analysis module such as described below, taking the words entered into the computer system as another source of input on which phrase matching and scoring can be performed.

10
With reference now to Fig. 10, there is illustrated a keystroke/mouse analysis module 59 that analyses the screen content and/or keystroke/mouse recordings that can be made at a CSR's PC. Three independent analyses of the keystrokes

15 provide for the following.

First, word and phrase identification 62 can be achieved by combining successive keystrokes into words and then phrases since the module can make the keystroke information a useful

20 search field. The module 59 must take account of the use of mouse clicks, tab keys, enter key etc. that delimit the inputs into a specific field and hence separate subsequent text from that entered prior to the delimiter. Interval analysis 64 is achieved by analyzing the time between

25 successive keystrokes.

Secondly, an indication of typing skills can be obtained. The use of specific keys such as backspace and delete can also give indications of level of typing accuracy. The

30 results of this analysis are useful in targeting typing training courses at those most likely to benefit.

Finally, a range analysis function 65 can be achieved by noting the variety of keys used and compared against other

35 calls. It is then possible to identify users who are

unfamiliar with, for example, standard windows shortcut keys (Alt+C) or application specific shortcuts (F2 for order form). The frequency of use of these less common keystrokes can be stored and subsequently used to identify

5 opportunities for windows and/or application specific training.

The outputs of the above stages may be accumulated at location 66 through the call and the net results stored in

10 addition to the individual instances.

The output of this module can again comprise database entries 68 for the call and/or file content 69 listing the results of the analyses 62,64,65,66 discussed above.

15

20

25

30

35

CLAIMS

1. A method of monitoring sets of related communication signal streams comprising the steps of analysing the content or parameters associated with a component of one of the signal streams according to a first analysis criteria;

   analysing a second component of a related signal stream or parameter associated therewith, according to a second analysis criteria;

   providing results of the analysis of the said one of the signal streams and which is responsive to the said analysis according to the second criteria.

2. A method as claimed in Claim 1, wherein the said first analysis criteria is selected by means of the said second criteria.

3. A method as claimed in Claim 1, wherein the said first analysis criteria is arranged to be adapted by means of the said second criteria.

4. A method as claimed in Claim 1, 2 or 3, and wherein the said analysis of the said content or parameters and the analysis of the signal stream are combined to provide a composite output parameter.

5. A method as claimed in any one or more of Claims 1-4, wherein the analysis according to the second criteria occurs prior to the analysis according to the said first criteria.

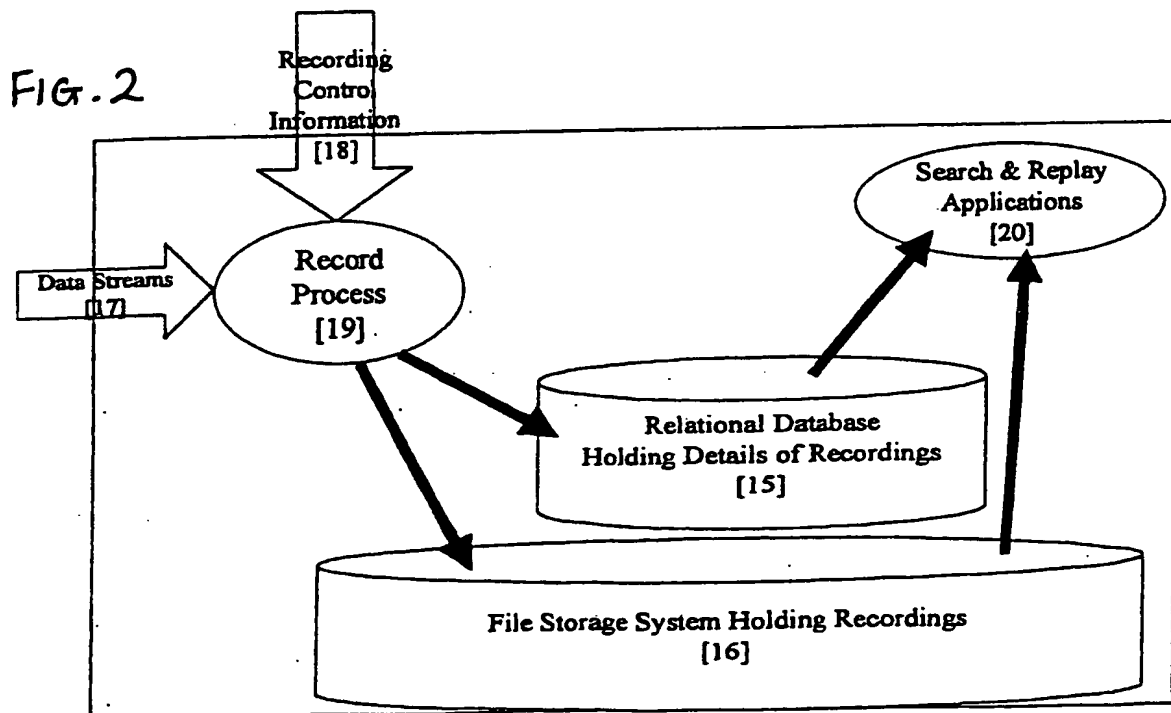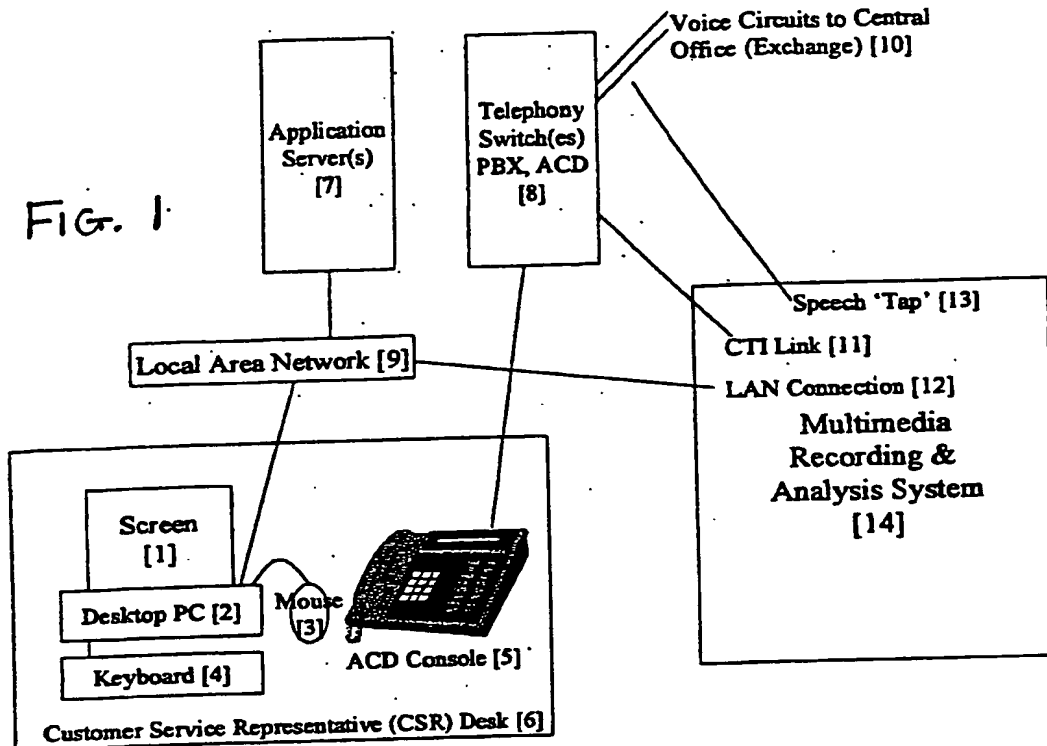6. A method as claimed in any one or more of Claims 1-5, and including the step of recording the signal stream.

7. A method as claimed in any one or more of Claims 1-6, and including the step of introducing timing information serving to locate analysed portions within the signal stream.

8. A communications monitoring system and including means for executing the method of any one or more of Claims 1-7.

9. A communication monitoring method including the steps of determining an energy envelope representative of at least one communication signal, and providing for the subsequent analysis of the said energy envelope.

10. A method as claimed in Claim 9, wherein at least two energy envelope files are employed.

11. A method as claimed in Claim 9 or 10, and arranged to allow for the selective analysis of the energy envelope.

12. A method as claimed in Claim 11, and arranged to allow for analysis of the energy envelope representative of the final section of the communication signal.

13. A method as claimed in Claim 9, 10, 11 or 12, and including the step of analysing the energy envelope so as to identify clipping of the signal.

14. A method as claimed in Claim 9, 10, 11, 12 or 13, and including the step of determining sound/silence ratios from the energy envelope.

15. A method as claimed in any one or more of Claims 9 to 14, and including the step of analysing the duration of sound passages.

16. A method as claimed in any one or more of Claims 9 to 15 and including the step of analysing the delays between signal transmissions in different directions.

17. A method as claimed in any one or more of Claims 9 to 16, and including the step of storing the energy envelope for analysis.

18. A method of monitoring a communication signal as defined in any one or more of Claims 1 to 7 and including the method steps of any one or more of Claims 9 to 17.

19. A communications monitoring system and including means for executing the method as defined in any one or more of Claims 9 to 17.

20. A communications monitoring method including the steps of conducting speech recognition for the identification of words and/or phrases within a communications traffic stream, and including the step of varying the scale and/or nature of recognition analysis applied for the speech recognition responsive to the analyses of content or parameters associated with the communications stream or related streams.

21. A method as claimed in Claim 20, wherein the scale and/or nature of the recognition analysis is arranged to be varied responsive to the identification of at least one party to the communication session.

22. A method as claimed in Claim 20, wherein the scale and/or nature of the recognition analysis is arranged to be varied on the basis of the length and/or stage of the communication session.

23. A method as claimed in Claim 20, 21 or 22, and including the step of generating a score signal indicative of such a level of satisfaction

24. A method as claimed in Claim 20, 21, 22 or 23, and including the step of monitoring the operation of a user interface device, the output of which is employed in controlling or adapting the recognition analysis.

25. A communication monitoring method of any one or more of Claims 1 to 7 and 9 to 18, and including the steps of Claims 20 to 24.

26. A communications monitoring system and including means for executing the method steps of any one or more of Claims 20 to 25.

27. A communications monitoring method including the step of monitoring usage of a user-interface device associated and arranged to be used concurrently, with the communication stream and controlling the communications monitoring responsive to the results of said monitored usage.

28. A method as claimed in Claim 27, and including the step of monitoring the accuracy with which a user employs the said interface device.

29. A method as claimed in Claim 27 or 28 and wherein the said user-interface device comprises a computer device.

30. A method as claimed in Claim 29 and including the step of monitoring the keystrokes and/or mouse actions of the user.

31. A method as claimed in Claim 29 or 30 and including the steps of monitoring the applications, documents and/or windows selected by the user.

32. A method as claimed in any one or more of Claims 27 to 31, and including the step of delineating different sections of a record of use of the said interface device so as to associate such different sections with respective different sections of the monitored communication.

33. A method as claimed in any one or more of Claims 27 to 32, and including the step of monitoring jointly the use of the said interface device and the level and/or nature of communications traffic to identify characteristics of the user.

34. A communications monitoring method of any one or more of Claims 1 to 7, 9 to 18, 20 to 24 and including the steps of any one or more of Claims 26 to 33.

35. A communications monitoring system including the means for executing the method steps of any or more of Claims 27 to 34.

FIG. 1

Application Server(s) [7]

Telephony Switch(es) PBX, ACD [8]

Voice Circuits to Central Office (Exchange) [10]

Local Area Network [9]

Speech 'Tap' [13]

CTI Link [11]

LAN Connection [12]

Multimedia Recording & Analysis System [14]

Screen [1]

Desktop PC [2]

Mouse [3]

Keyboard [4]

ACD Console [5]

Customer Service Representative (CSR) Desk [6]

FIG. 2

Recording Control Information [18]

Data Streams [17]

Record Process [19]

Search & Replay Applications [20]

Relational Database Holding Details of Recordings [15]

File Storage System Holding Recordings [16]

FIG. 3

Recording
Control
Information
[18]

Data Streams
[17]

Record
Process
[19]

Search & Replay
Applications

Relational Database
Holding Details of Recordings
[15]

File Storage System Holding Recordings
[16]

FIG. 4

Recording
Details In
[25]

Recording
Content In
[26]

Analysis Module
[24]

Derived Recording
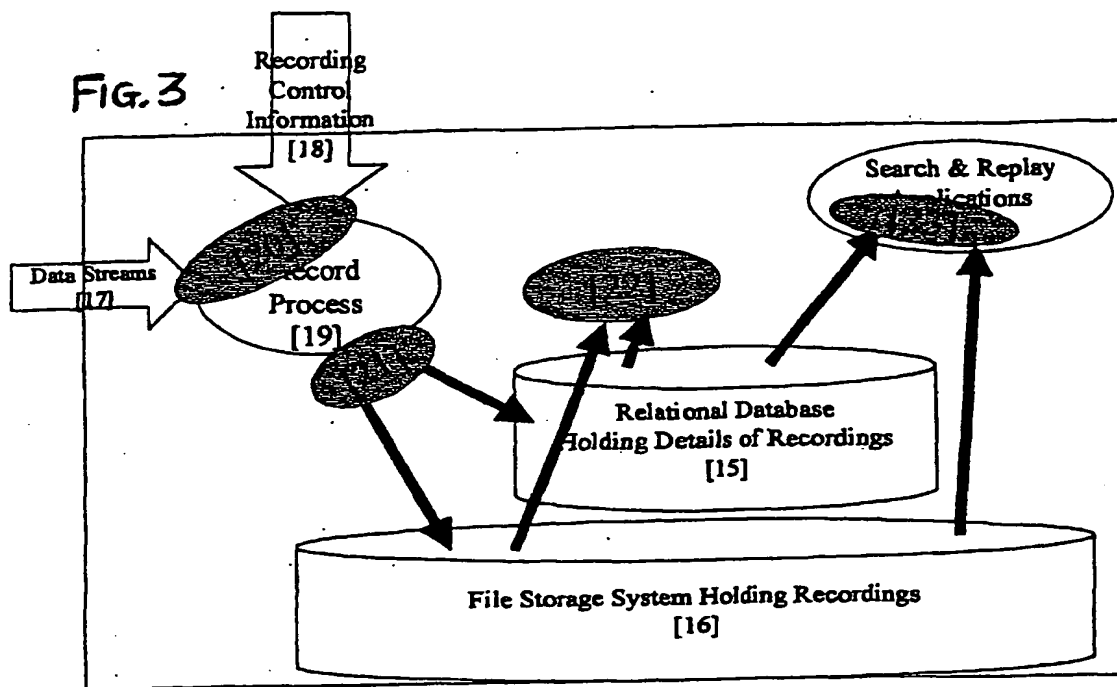Details Out
[27]

Derived Recording
Content Out
[28]

BEST AVAILABLE COPY

FIG. 5
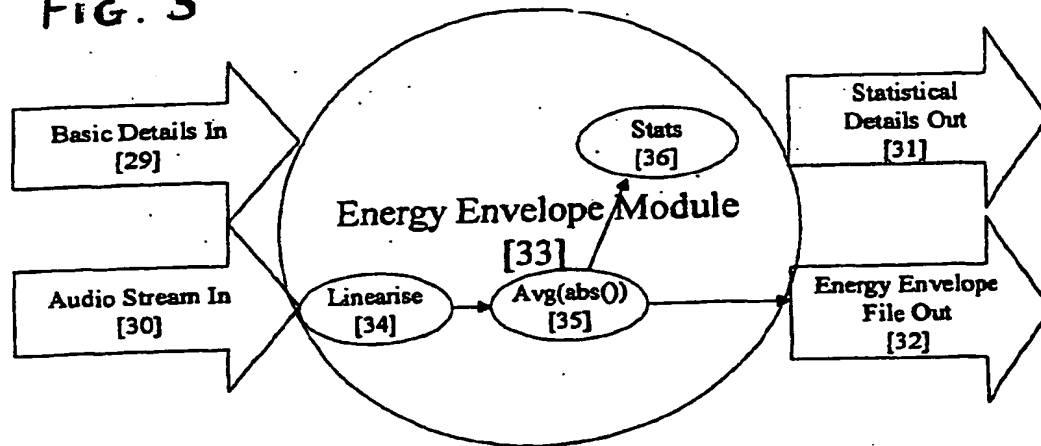

FIG. 6A

FIG. 6B

FIG. 7

Existing Call Details In [40]

Energy Envelope Files In [38]

Energy Envelope Analysis Module [37]

Stereo Pair Compare [39]

Stats [41]

Derived Call Details Out [42]

FIG. 8

Basic Details In [45]

Audio Stream In + Energy Envelopes [44]

Speech Recognition Module [43]

Call 'slicer' [46]

Recognition Engine [47]

Database Uploader [48]

Transcript Details Out to Dbase [49]

Transcript File Out [50]

FIG. 9

Basic Details In [52]

Transcript + Word List In [53]

Language Analysis Module [51]

Phrase Reco [54]

Database Uploader [56]

Score Accumulator [55]

Phrase Matches + Stats Out to Dbase [57]

Phrase File Out [58]

FIG. 10

Basic Details In [60]

Screen Content Keyboard/Mouse [61]

Keystroke/Mouse Analysis Module [59]

Phrase Reco [62]

Interval Analysis [64]

Range Analysis [65]

Score Accumulator [66]

Database Uploader [67]

String Matches + Stats Out to Dbase [68]

Strings File Out [69]